DOCUMENT RESUME

ED 216 332 CS 006 673

AUTHOR Otto, Beverly; Sulzby, Elizabeth
TITLE Judging the Emergent Reading Abilities of Kindergarten Children.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Dec 81
GRANT NIE-G-80-0176
NOTE 21p.; Paper presented at the Annual Meeting of the National Reading Conference (31st, Dallas, TX, December 2-5, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Beginning Reading; Behavior Rating Scales; Kindergarten Children; *Measurement Techniques; *Measures (Individuals); Primary Education; *Reading Ability; *Story Reading; *Student Evaluation; Writing Skills

ABSTRACT

In 1981, a scale, the Emergent Reading Ability Judgments for Dictated and Handwritten Stories, was developed for use in assessing how close a child was to reading independently based upon the nature of the child's attempts to read from dictated and handwritten stories. A study was conducted to apply the scale to stories from a new sample of children whose overall emergent reading abilities were ascertained by other techniques. The question addressed was whether the scale covers the reading attempts of these children. Subjects for the study were nine kindergarten children who fell into high, moderate, or low emergent reading ability groupings. The children's own dictated and handwritten stories comprised the materials used for reading attempts. Findings showed that the scale could be applied to repeated reading attempts of kindergarten children. Further analysis indicated that the scale discriminated between groups of children who differed in overall emergent abilities. There was also a trend in the data for children to improve across time. In addition, the scale showed moderate and, in two of three instances, significant ability to rank subjects similarly across dictated and handwritten forms within a given trial. Finally, even though the two measures were devised upon different bases for different purposes, the scale and a traditional readiness test also tended to rank children similarly. (HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED216332

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☑ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Judging the Emergent Reading Abilities of Kindergarten Children

Beverly Otto

and

Elizabeth Sulzby*

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Beverly Otto
Elizabeth Sulzby
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Northwestern University
Evanston, Illinois

Presented at the National Reading Conference

Dallas, Texas
December 1981

---

* The research reported in this paper was supported in part by a grant from the National Institute of Education, NIE-G-80-0176, to the second author. The statements made are solely the responsibility of the authors.

CS006673

Dr. Elizabeth Sulzby
Assistant Professor
School of Education
Northwestern University
2003 Sheridan Rd.
Evanston, Illinois 60201
(312) 492-3620

Judging the Emergent Reading Abilities of Kindergarten Children[1]

Beverly J. Otto
Northwestern University

Elizabeth Sulzby[2]
Northwestern University

The purpose of this paper is to describe the development and validation of an assessment tool for judging the emergent reading abilities of young children. In particular, this tool is designed for use with children's reading attempts using their own stories, both handwritten and dictated, as the texts. (The term "story" is used here to indicate the composed text, regardless of any genre classification, because the term "story" was used with the children.)

The term "emergent reading ability," as described by Holdaway (1979), emphasizes the importance of early reading-like behavior as being true literacy skills, not merely indicators of reading readiness in the traditional sense of predicting achievement in a yet-to-be-acquired skill. Emergent reading abilities can be considered to be readiness in the sense, however, of providing readiness for the next step in on-going acquisition.

Researchers who choose to investigate the transition period from pre-reading or emergent reading to independent reading are faced with some unique problems, including how to describe the nature of reading materials for beginners. The position being taken in this study is that any tool for judging emergent reading attempts must be based upon an analysis of the fit between the text being "read" and the knowledge of the young "reader" about that text. Thus, a tool that fits the child's attempts to read his/her own dictated and handwritten texts will have characteristics specific to those kinds of texts. It may be similar to tools used for emergent reading of other text types like storybooks (Haussler, 1981; Sulzby, 1981b) or basal readers (Clay, 1978) but it will be constrained by the intimate relation between reader and his/her own text.

Various analyses have been made between text types and early reading. Sulzby (1981b) contends that both handwritten and dictated stories support emergent comprehension as exhibited in memory for text but do not tend to support comprehension as construction of intended

meaning. Dictated stories tend to support memory for conventionally-spelled units while handwritten stories tend to give more support to use of letter-sound information but not necessarily in conventional spelling form. Holdaway (1979) argues that the child's own compositions consist of less memorable language than do favorite storybooks; thus storybooks support memory for text more than do child compositions, in his view. Neither researcher has yet published a comparative study between the text types nor does this study propose to make such a comparison. Instead, it addresses the question of whether a scale for the child's dictated and handwritten compositions can produce reliable and valid estimates of emergent reading for those texts.

Dictated stories have been described as important first texts for young readers. Most of the arguments supporting these texts have consisted of illustrative examples and/or rationales. Stauffer (1970) suggested that using children's oral language "written down" in dictation is an appropriate technique for inducting a child into independent reading. The advantage of this approach involves the fit between the text and the knowledge of the young reader; difficulty level is controlled because the syntactic patterns, topics, and vocabulary choice come from the child him/herself.

Support for the use of the child's own written compositions is somewhat indirect. Primarily, it comes from descriptive studies that indicate that, while reading and writing may not be considered to be mirror-images (Bissex, 1980), they do develop within the same time frame

and appear to be importantly related (Gentry & Henderson, 1978; Read, 1971). Chomsky (1971) proposes that the child should be encouraged to "write first, read later" because the child can learn much about reading from writing. Graves (1979) and Sulzby (1981b) indicate that the child engages in much re-reading as a voluntary part of composition. Such composition includes but is not restricted to the use of invented spelling (Clay, 1978; Sulzby, 1981b).

Sulzby (1981c) reported the development of a scale for use in assessing how close a child is to reading independently based upon the nature of the child's attempts to read from dictated and handwritten stories. In dictation, a scribe writes the words which a child says in response to a "dictate" request. In handwritten composition, the child uses whatever writing system s/he prefers. The scale was developed from both the nature of the composition and the re-reading attempt.

Sulzby applied the scale to pairs of dictated and handwritten story-reading attempts of 24 kindergarten children; the study included a complete replication of story sets, leading to 96 judgments, 48 each for dictated and handwritten stories. Agreement between judges was high (96%) and the rank order correlation between the two replications each containing 48 attempts was rho = .77 (corrected for ties), $p<.01$. She reported a rank order correlation between average scores for four applications of the scale and the Metropolitan Readiness Tests, Form P, Level 1, of rho = .37 (corrected for ties), $p<.05$. She interpreted these scores to indicate that the scale did have acceptable reliability

in scoring, consistency across time with the same children, and some relationship with scores derived from traditional measures for predicting reading achievement (readiness tests).

The Sulzby (1981c) study was limited in scope to the consistency of the scale and to the relationship between the scale and a traditional readiness test. There was no convincing assessment of the scale as a valid measure of emergent reading abilities of any kind besides the reading of dictated and handwritten stories. Additionally, the scale was applied to the data from which it was derived so questions remain about whether there might be significant categories of behavior that were either not exhibited or not detected in the original sample.

These limitations and the need to further explore the emergent reading abilities of kindergarten children led to the current study. This study was designed to address more specific issues of the reliability and validity of the Emergent Reading Ability Judgments for Dictated and Handwritten Stories (Sulzby, 1981c). The current study applies the scale to stories from a new sample of children whose overall emergent reading abilities were ascertained by other techniques.

This study addresses the question of whether the scale covers the reading attempts of these children. If it covers the behaviors (a validity issue), can trained judges apply the scale with sufficient agreement and how reliable are repeated administrations of the scale to reading attempts of the same children across trials?

In addition to the ability of the scale to cover the new data, other validity issues include the following concerns. How well does it discriminate subject groups chosen on overall emergent reading ability? How well does it discriminate subject groups' achievement across time? How well does it describe behavioral change across time? Finally, how does it compare with a test of reading readiness?

Method

Subjects

Nine children were selected for case studies from among 24 children in a kindergarten classroom in an upper middleclass suburb north of Chicago, Illinois. The children were identified as being either high, moderate, or low in overall emergent reading ability at the beginning of the kindergarten year of 1980-81. This identification was based upon results of an interview study, "General Knowledges About Written Language," with teacher judgments used to refine the researchers' selections. Full procedures may be found in Sulzby (1981a).

The case study sample consisted of five girls and four boys, ranging in age from 4-11 to 5-11 in October. Three children fell into each group, with sex distribution not equal within groups. The high group consisted of three girls; middle group, of two girls and one boy; and the low group, of two boys and one girl. This distribution was reflective of the distribution of overall emergent reading ability within the 24-child classroom.

Data

The children's own compositions comprised the materials used for reading attempts. Dictated and handwritten story pairs were collected during the course of two different studies. Study II, "Real and Make-Believe Topic," contributed two complete sets, stories about learning to ride a big wheel. Study IV, "New Event Stimulus," yielded one set of stories about a race with novelty wind-up toys that triads of children participated in just prior to composition and re-reading. For purposes of this paper, those three story sets will be referred to as Trial 1, Trial 2, and Trial 3. Trial 1 took place in January and early February; Trial 2 took place in February; and Trial 3, in April and May.

Scoring

Two trained examiners scored the full collection of story/reading attempts. Initial scoring was done independently, using the Sulzby (1981c) scale with conferences over disagreements. There were 54 story/reading pairs, with 27 dictated and 27 handwritten attempts. Judges were able to come to complete agreement but it should be noted that the agreement was conferred rather than independent for reasons that follow.

Initial scoring yielded 12 cases that did not fit the scale. These cases were of three kinds: one time the child produced what may be called a "pattern" story rather than the story requested; two times an examiner had given an illegitimate prompt during the re-reading attempts; and nine times the child's behavior did not fit the scale as originally worded. This final group consisted of one kind of behavior that was

judged to be the equivalent of Sulzby's "five" category. The initial "five" category was defined as an instance in which the child's eyes were not on print but his/her story rendition was stable; the instances in which the scale did not fit were all times when the child's eyes were on print (not tracking) but his/her story was not stable. These behaviors were clearly more advanced than "four," and not so advanced as "six" in which the child is able to maintain eyes on print and a stable story, without actually reading the print independently. Table 1 gives the scale with this revision and with wording modifications for clarity.

Insert Table 1 about here or at end

(Editor's preference)

After the scale was revised the scorers were easily able to classify 89% of the attempts, with the exception of the other three mentioned above. These three cases were discussed until the examiners agreed upon a placement for the attempts based upon inferences from surrounding evidence such as storytelling attempts, samples of aspects of reading, and other comments from the full protocols. One of these judgments resulted in a data point that seemed to be outside the general developmental pattern but the placement seemed to be consistent with the child's performance within the bounds of the scale; this was the "pattern" story in which the child dictated series of "I like my

mom; I like my dad; I like my cat," sentences rather than the big wheel story.

Analysis

To examine the ability of the scale to distinguish between groups of children judged to be of differing ability, a 3 (Groups: high, moderate, low) X 2 (Mode: dictated, handwritten) X 3 (Trials: 1, 2, 3) mixed analysis of variance was conducted with modes and trials as repeated factors. Rank order correlations were calculated for each dictated/handwritten trial and for comparing the set of judgments with results of the Metropolitan Readiness Test, Level I, Form P (11/4/80).

Results

Results of the analysis of variance are shown in Table 2. These findings indicate that the scale differentiated between groups significantly, $F(2,6)=9.03$, $p<.025$. The means for the groups were as follows: High, 6.2; Middle, 4.4; and Low, 3.3. Two of the three children in the high group were at ceiling levels (total scores for six observations were 42 and 41 or 42), thus the differences between groups can be considered to be conservative.

Insert Table 2 about here

In addition to group differences, the only other significant result of the analysis of variance was the interaction between Mode and

Trials, $F(4,32)=5.31$, $p<.01$. Figure 1 shows that interaction graphed; dictated story reading attempts tended to be scored slightly higher than handwritten stories for the final two trials, but for Trial 1 they were quite a bit higher (mean of 5.2 compared with 3.4).

Insert Figure 1 about here

Another question about group differences can be addressed by examination of the means for the nonsignificant Groups by Trials interaction. The question is two-fold: (1) were groups distinguished across time, and (2) did children increase in Emergent Reading Abilities for Dictated and Handwritten Stories over time? While the main effect for groups answered the first part of the question in the affirmative, the means displayed in Table 3 show that, while the high group was almost at ceiling and stayed stable over time, the low and middle group mean scores increased across trials.

Insert Table 3 about here

The rank order correlations (corrected for ties) between the pairs of dictated and handwritten stories were as follows: Trial 1, rho=.83 ($p<.01$); Trial 2, rho=.58 (NS); and Trial 3, rho=.73 ($p<.05$). The correlation between the total scores for the ERAJ for the six ses-

sions and the total raw scores on the Metropolitan Readiness Test, Level 1, was rho=.80 ($p<.05$, uncorrected since there was only one tie).

These correlations indicate that the rank order of subjects between the paired administrations of the scale tended to be maintained. Additionally, the rank associated with the subjects' total scores for the six sessions tended to be close to the rank of subjects' scores from the Metropolitan.

Discussion

The scale, Emergent Reading Ability Judgments for Dictated and Handwritten Stories, was able to be applied to repeated reading attempts of kindergarten children. This application required one revision in order to capture behavior not covered by a previous version of the scale. Revision of the scale resulted in an assessment tool which could be used by trained examiners with high agreement.

The scale is limited to children's reading attempts using their own compositions, dictated and handwritten, as the texts. A further limitation of the scale is that people using it must be knowledgeable about the nature of young children's writing systems. The scale requires the assessment both of the composition and of the re-reading attempt. It is thus a one-to-one assessment tool. Finally, the scale only makes gross discriminations of ordinal level. It does not give finer diagnosis of the child's attempts as the child nears independent reading. Thus it may imply a more rigid sequence than a closer examination of exact reading-like behaviors reveal.

Nevertheless, the analysis of variance over three repeated administrations of the scale to dictated and handwritten story reading attempts indicated that the scale discriminated between groups of children who differ in overall emergent reading abilities. While the main effect for trials was not significant, there was also a trend in the data for children to improve across time, particularly for handwritten stories.

In addition to the ability of the scale to discriminate groups of children the scale also showed moderate and, in two of three instances, significant ability to rank subjects similarly across dictated and handwritten forms within a given trial. This correlation is worth noting, given the mode by trial interaction (see Figure 1). While the dictated mean for Trial 1 was significantly higher than the handwritten mean, the ranks of scores for those two administrations were significantly related.

Finally, even though the two measures are devised upon different bases for different purposes, the scale and a traditional readiness test also tended to rank children similarly. The correlation for this comparison was the children's total ERAJ scores across four to five months and the total raw score for the readiness test. It would be surprising and somewhat suspicious if the two measures were totally unrelated since both purport to measure some facets associated with written language.

It should be remembered that the correlations reported are for ranks and are for nine children. Hence, the significance level is more important than the magnitude. Sulzby (1981c) reported significant rank order correlations for a different and larger sample (N=24).

The Emergent Reading Ability Judgments for Dictated and Handwritten Stories has two potential uses. It can be used, along with other tools, in research in early literacy development. It can also be used by teachers of young children to provide one measure of ongoing emergent reading and writing abilities. In both kinds of use, it is limited to being a gross measure of structural characteristics. It needs further refinement and should be used in combination with other assessment tools; nevertheless, it seems to provide some assurance of being one reliable and valid assessment tool for reading behaviors of young children who are not yet reading independently.

Footnotes

[1]The research reported in this paper was supported in part by a grant from the National Institute of Education, NIE-G-80-0176, to the second author. The statements made are solely the responsibility of the authors.

[2]The authors thank the children and their teacher who must remain anonymous. We also thank Susan Anderson, Beverly Cox, and Dr. Norman Bowers for their assistance.

References

Bissex, G. L. GNYS AT WRK: A child learns to write and read. Cambridge: Harvard University Press, 1980.

Chomsky, C. Write first, read later. Childhood Education, 1971, 47, 296-299.

Clay, M. M. Reading: The patterning of complex behavior. 2nd ed. Auckland: Heinemann Educational Books, 1978.

Gentry, J. R., & Henderson, E. H. Three steps to teaching beginning readers to spell. The Reading Teacher, 1978, 31, 632-637.

Graves, D. H. The growth and development of first grade writers. Paper presented at the annual meeting of the Canadian Council of Teachers of English. Ottawa, May 1979.

Haussler, M. Print awareness and its role in beginning reading. Paper given at the 26th Annual Convention of the International Reading Association, New Orleans, April 1981.

Holdaway, D. The foundations of literacy. Sydney: Ashton Scholastic, 1979.

Read, C. Preschool children's knowledge of English phonology. Harvard Education Review, 1971, 41, 1-34.

Stauffer, R. G. The language-experience approach to the teaching of reading. New York: Harper and Row, 1970.

Sulzby, E. Beginning readers' developing knowledge about written language. Interim report to the National Institute of Education. Evanston, IL: Northwestern University, 1981. (a)

Sulzby, E. Kindergarteners begin to read their own compositions: Beginning readers' developing knowledges about written language project. Final report to the Research Foundation of the National Council of Teachers of English. Evanston, IL: Northwestern University, 1981. (b)

Sulzby, E. When kindergarteners read and write: Told and dictated story characteristics. Paper presented at the Annual Meeting of the American Educational Research Association, Los Angeles, April 1981. (c)

Table 1

Emergent Reading Ability Judgments
for Dictated and Handwritten Stories, Revised

Score assigned	Behaviors observed
(For 1-2, any production cannot be called a story or composition either by intention or by characteristics of the product.)	
1	No dictated nor handwritten stories; hence, no attempt to re-read. Child refuses to pretend-write.
2	In response to request to handwrite a story, no "story" is produced but there is some primitive evidence of reading and writing. In response to request to dictate a story, child says something that does not seem to be treated as a composed unit; it may consist of conversational characteristics either in wording or in turn-taking. It is either not re-read or re-reading attempt is very primitive.
(For 3-7, stories or their equivalent are produced.)	
3	While stories are produced, child refuses to re-read. For own writing, child may say that the story "doesn't say anything," or, for dictation, "I can't read."
4	Child attempts to re-read but eyes are not on print. Story recited is similar to original but not stable.*
5	This seems to be a period during which the child cannot simultaneously maintain eyes on print and recite stable story. Score "5" either if (a) child's eyes are on print but story recited is similar but not stable* or (2) child's eyes are not on print but story recited is stable.*
6	Eyes are on print but child is not tracking print. Story recited is stable.* For handwritten story, print may still be pretend-cursive or other writing system <u>if</u> the story accompanied the composition, either as shown by voicing during composition or by other evidence such as mixture of invention or reference to units by intention.
7	Child's eyes are tracking print, matching voice to print, "actually reading," independently, with attention to meaning. To be rated as a "7" story must consist of more than one complete clause; if less, then rank "6". Print can have mature forms of invented spelling or of combinations of other writing units as long as child treats individual units as stable.

*Stable in reference to re-reading stories means that no clause-level units have been added, omitted, or placed out of composed sequence.

Table 2

Emergent Reading Ability Judgments (ERAJ) for Dictated and Handwritten Stories of Children of Differing Abilities Across Composition Mode and Trials

Source	SS	df	ms	F	p
Total	151.31	53	---	---	
Between	102.48	8	---	---	
Groups (G)	76.92	2	38.46	9.03	<.025
$Error_b$	25.56	6	4.26	---	
Within	51.83	45	---	---	
Mode (M)	8.16	1	8.16	4.61	NS
Trials (T)	6.03	2	3.01	1.45	NS
G x M	2.12	2	1.06	< 1	NS
G x T	4.19	2	2.10	1.01	NS
M x T	6.79	4	1.70	5.31	<.01
G x M x T	2.77	4	.69	2.16	NS
$Error_1$	3.55	2	1.77	---	
$Error_2$	8.11	4	2.08	---	
$Error_3$	10.11	32	.32	---	

Table 3

Mean Emergent Reading Ability Judgments for Dictated and Handwritten Stories: Groups by Trials

Groups	Trial 1	Trial 2	Trial 3
High	6.3	6.2	6.2
Middle	3.8	4.2	5.2
Low	2.8	3.2	4.0

Figure 1

Mean Emergent Reading Ability Judgments for Dictated and Handwritten Stories: Mode by Trials